TGT Training Loss with Random Embeddings

Loss (log)

Epochs

— One Layer Encoder    — Two Layer Encoder